

REMARKS

In the application, claims 1-21 are pending and rejected. After due consideration of the Examiner's comments in the Office Action of February 17, 2005, the claims have been amended to more clearly set forth what Applicants regard as their invention. Applicants respectfully request reconsideration of the claims as amended.

Rejections under §112

Claims 1-21 are rejected under 35 U.S.C. §112, 2nd para., as being vague and indefinite.

With regard to claim 1, the Examiner states that there are no steps or acts of "managing" and/or "retrieving" biological data as indicated in the preamble.

The step of acquiring and staging gene expression data for storage has been added, however, it is respectfully submitted that the word "managing" cannot be taken out of context and that the step of "providing a data warehouse" is a key component of managing the data for efficient exploration and analysis. It is the management of the data storage that facilitates access for searching the data. As clearly described in the specification, the invention is directed, in part, to a method of managing gene expression and related data. See, e.g., page 6, line 1, "The present invention uses data warehousing methodology to *manage* and explore gene expression and related data."; page 6, line 13, and "Data warehouse *management* tools are used for maintaining data consistency...." (*emphasis added.*)

The "retrieving" aspect must also be taken in the context of the preamble, meaning that the full act is the retrieving of data for efficient exploration and analysis. It is submitted that this act is the combination of the remaining steps of the base claims: providing a user interface that permits entry of a query, receiving the query, correlating gene expression measurements with the query, and displaying the results at the user interface. Support is provided throughout the specification for the act of retrieving data. See, e.g., page 29, line 20, "The present invention pertains to relational databases for storing and *retrieving* biological information comprising an integration of at least three databases organized to support exploration and mining of gene expression data."; and page 30, line 20, "The key objective of the relational database for storing and *retrieving* biological information of the present invention is to provide comprehensive access to gene expression and support for biological analysis."

With regard to claims 1, 8 and 15, the Examiner finds that the term “first” is confusing without a “second” term.

It is submitted that the adjective “first” is used in the base claims to avoid confusion when introducing the different, “second sample set or gene set” in claim 5.

With regard to claims 1, 8 and 15, the Examiner finds the language “corresponding to” as vague and indefinite. The claims have been amended to clarify the relationship between the tissue and cell lines and the sample data and DNA fragments. It is submitted that it would, nonetheless, be readily apparent to one of skill in the art that there are relationships between the data in the three databases, and that sample data corresponds to gene expression data, as does the DNA fragment data.

With regard to claims 1 and 15, the Examiner finds the language “responsive to” as vague and indefinite.

That claim language has been amended to further clarify the relationship between the gene expression measurements and the query, specifically that what is displayed at the user interface are the gene expression measurements that are responsive to the query.

With regard to claims 4, 5, 11, 12, 18, and 19, the Examiner finds the language “consistently expressed” & “consistently not expressed” as vague and indefinite.

The claims have been amended to replace “consistently expressed” with the word “present” and “consistently not expressed” with the word “absent”.

With regard to claims 4-7, the Examiner finds the limitation “wherein the analysis comprises...” lacks proper antecedent basis. Claims 4-7 have been amended to eliminate the term “analysis”.

Rejection under §102

Claims 1, 4, 5, 8, 11, 12, 15, 18, and 19 are rejected under 35 U.S.C. §102(b) as being anticipated by Bassett et al. (*Nature Genetics Suppl.*, V. 21, pp. 51-55, Jan. 1999).

The Examiner responds to Applicants’ prior assertion that Bassett et al. fails to teach the partitioning of a data warehouse by referring to Bassett et al.’s proposal for construction of a data warehouse where the essential information can be assigned to different categories, including gene expression data, clinical or sample data and a fragment index. The Examiner appears to

allege that a data category, i.e., a type of data, is the same as a database. Applicants' respectfully disagree.

While Bassett et al. teach a method for data management and analysis of gene expression data, they do not disclose any detail of their database structure, much less teach the use of separate databases for separately storing different categories of data as claimed by Applicant. In Bassett et al.'s discussion of data warehousing beginning on page 53, they write that "[t]he first step is to construct a database or 'data warehouse.'" They are equating database and data warehouse. Nowhere in the article do Bassett et al. refer to or suggest multiple databases within a data warehouse. There is a single database, which according to their usage is the same as a data warehouse.

The definition of a data warehouse available from "The Data Warehousing Information Center" on the World Wide Web at dwinfocenter.org is "a copy of transaction data specifically structured for querying and reporting." "The form of stored data has nothing to do with whether something is a data warehouse. A data warehouse can be ... a relational database, multidimensional database, flat file, hierarchical database, object database, etc.." The body of the text of Bassett et al. does not describe the form of their database, however, Fig. 1, an overview of the information system for large-scale gene expression experiments, specifies the database as a relational database, citing to Ermolaeva et al. ("Data management and analysis for gene expression arrays", *Nature Genet.* 20:19-23 (1998), a copy of which is attached as Exhibit A). Definitions of "relational database" are provided in attached Exhibit B, which includes an article entitled "Introduction to Relational Databases", published June 24, 2002 in *Database Journal*, available on the World Wide Web at databasejournal.com, and printouts from two Internet dictionaries: dictionary.net and dict.die.net. Relevant portions are highlighted. Each reference provides the same definition, that a relational database is a collection of tables. Tables are collections of records arranged in rows and columns. The database described by Ermolaeva et al. is "an industry standard relational database management system" (page 20, 2nd column, line 6). Included with the Ermolaeva et al. article in Exhibit A is a copy of the layout of the relational schema of the database, which is available from the authors on request. (See the last sentence in the caption of Fig. 2.) It is readily apparent from this layout that different data categories such as gene expression data, experiment data and fragment data are represented as different tables

within a single relational database. For example, the table labeled “Intensities” is the same category of data that would be stored in Applicants’ gene expression database, while the “Unigene” and “Sequences” tables contain the same type of data that would be stored in Applicants’ Fragment Index database. A table is not a database, but only part of a database. Ermolaeva, et al. and Bassett et al. by their reference to Ermolaeva et al., teach only a single relational database, not a data warehouse with multiple distinct databases as disclosed and claimed by Applicants.

In addition to the above comments, it is submitted that Bassett et al. do not teach staging of the various data types for providing links between the separate databases. Because Bassett et al. teach the use of a single database, presumably a relational database, there is no staging database needed to provide a link between multiple databases as claimed by Applicants.

The partitioning of the data warehouse in three distinct databases as claimed by Applicants provides the advantage of greater accessibility for maintaining and updating data as scientific methods evolve. Further, the separate databases can be explored collectively or independently, thus providing faster access to the desired data. Bassett et al. fails to teach the partitioning of a data warehouse into three separate databases, where each database is dedicated to a different type of data. Further, Bassett et al. do not teach a staging database for providing links between the data in the different databases. Accordingly, Bassett et al. cannot anticipate Applicants’ invention as claimed, and Applicants respectfully request that the Examiner withdraw the rejection under §102.

Rejection under § 103

Claims 1-5, 8-12, and 15-19 are rejected under 35 U.S.C. §103(a) as being unpatentable over Bassett et al. in view of Gopalkrishnan et al. Gopalkrishnan et al. are cited for their disclosure of star/snowflake schema in a data warehouse.

It is respectfully submitted that Gopalkrishnan et al. do not impart to Bassett et al. the missing elements of Applicants’ claimed invention that would be necessary to render Applicants’ invention obvious. Specifically, Gopalkrishnan et al. do not teach a data warehouse consisting of multiple distinct databases. In fact, Figure 1 of Gopalkrishnan et al. illustrate their data warehouse as being the consolidation of data from three separate data sources into a single integrated database, the opposite of what Applicants’ invention does.